

# Universal variance bounds for the Pearson family \*

G. Afendras<sup>†</sup>

Department of Mathematics, Section of Statistics and O.R., University of Athens,  
Panepistemiopolis, 157 84 Athens, Greece.

## Abstract

We use some properties of orthogonal polynomials to provide a class of upper/lower variance bounds for a function  $g(X)$  of an absolutely continuous random variable  $X$ , in terms of the derivatives of  $g$  up to some order. The new bounds are better than the existing ones.

MSC: Primary 60E15.

Key words and phrases: Completeness; Derivatives of higher order; Fourier coefficients; Orthogonal Polynomials; Parseval identity; Pearson family of distributions; Rodrigues-type formula; Variance Bounds.

## 1 Introduction

Let  $Z$  be a standard normal random variable and  $g : \mathbb{R} \rightarrow \mathbb{R}$  any absolutely continuous function with derivative  $g'$ . Chernoff, [8], using Hermite polynomials, proved that (see also the previous papers by Nash, [13] and Brascamp and Lieb [3])

$$\text{Var } g(Z) \leq \mathbb{E}(g'(Z))^2,$$

provided that  $\mathbb{E}(g'(Z))^2 < \infty$ , where the equality holds if and only if  $g$  is a polynomial of degree at most one – a linear function. This inequality plays an important role in the isoperimetric problem and has been extended and generalized by several authors; see, e.g., [7, 17, 5, 15, 10, 14] and references therein.

**Definition 1.1** (Integrated Pearson Family) *Let  $X$  be a random variable with density  $f$  and finite mean  $\mu = \mathbb{E}X$ . We say that  $X$  (or its density) belongs to the integrated Pearson family (or integrated Pearson system) if there exists a quadratic polynomial  $q(x) = \delta x^2 + \beta x + \gamma$  (with  $\delta, \beta, \gamma \in \mathbb{R}$ ,  $|\delta| + |\beta| + |\gamma| > 0$ ) such that*

$$\int_{-\infty}^x (\mu - t)f(t)dt = q(x)f(x) \quad \text{for all } x \in \mathbb{R}. \quad (1.1)$$

*This fact will be denoted by  $X$  or  $f \sim \text{IP}(\mu; q)$  or, more explicitly,  $X$  or  $f \sim \text{IP}(\mu; \delta, \beta, \gamma)$ .*

Let a random variable  $X \sim \text{IP}(\mu; q)$  and consider a suitable function  $g$ . Johnson, [11], established Poincaré-type upper/lower bounds for the variance of  $g(X)$  of the form

$$(-1)^n (\text{Var } g(X) - S_n) \geq 0, \quad \text{where} \quad S_n = \sum_{k=1}^n (-1)^{k-1} \frac{\mathbb{E}q^k(X)(g^{(k)}(X))^2}{k! \prod_{j=0}^{k-1} (1 - j\delta)}. \quad (1.2)$$

In particular, for the normal see [10, 15] and for the gamma see [15].

\*Work partially supported by the University of Athens' Research fund under Grant 70/4/5637.

<sup>†</sup>e-mail address: g\_afendras@math.uoa.gr

Afendras et al., [1], using Bessel's inequality, showed that

$$\text{Var } g(X) \geq \sum_{k=1}^n \frac{\mathbb{E}^2 q^k(X) g^{(k)}(X)}{k! \mathbb{E} q^k(X) \prod_{j=k-1}^{2k-2} (1-j\delta)}; \quad (1.3)$$

for the case  $n = 1$  see [4].

Afendras and Papadatos, [2], showed that, under appropriate conditions, the following two forms of Chernoff-type upper bounds of the variance of  $g(X)$  are valid:

$$S_{n,(\text{str})} = \sum_{i=1}^n \frac{\mathbb{E}^2 q^i(X) g^{(i)}(X)}{i! \mathbb{E} q^i(X) \prod_{j=i-1}^{2i-2} (1-j\delta)} + \frac{\mathbb{E} q^n(X) (g^{(n)}(X))^2 - \frac{\mathbb{E}^2 q^n(X) g^{(n)}(X)}{\mathbb{E} q^n(X)}}{(n+1)! \prod_{j=n}^{2n-1} (1-j\delta)}, \quad (1.4)$$

$$S_{n,(\text{weak})} = \sum_{i=1}^{n-1} \frac{\mathbb{E}^2 q^i(X) g^{(i)}(X)}{i! \mathbb{E} q^i(X) \prod_{j=i-1}^{2i-2} (1-j\delta)} + \frac{\mathbb{E} q^n(X) (g^{(n)}(X))^2}{n! \prod_{j=n-1}^{2n-2} (1-j\delta)}. \quad (1.5)$$

The equality in (1.4) holds when  $g$  is a polynomial of degree at most  $n+1$  and in (1.5) when  $g$  is a polynomial of degree at most  $n$ . The bound (1.5) for beta distributions is also given by [18].

In Section 2 we provide a new class of upper/lower bounds for the variance of  $g(X)$ . They can be called as ‘‘Poincaré-type’’ of order  $n$  and with point balance  $m$ . They hold for a subfamily of Pearson distributions. In particular, the bound for  $N(\mu, \sigma^2)$  distribution, namely

$$S_{m,n}(g) = \sum_{i=1}^m \frac{\binom{m}{i} \sigma^{2i}}{(m+n)_i} \mathbb{E}^2 g^{(i)}(X) + \sum_{i=1}^n (-1)^{i-1} \frac{\binom{n}{i} \sigma^{2i}}{(m+n)_i} \mathbb{E} (g^{(i)}(X))^2$$

(for  $(x)_k$  see Definition 2.1), satisfies the inequality

$$(-1)^n (\text{Var } g(X) - S_{m,n}(g)) \geq 0,$$

where the equality holds if and only if  $g$  is a polynomial of degree at most  $m+n$ .

For  $n$  fixed, Section 3 investigates the bounds  $S_{m,n}(g)$  as  $m$  increases. It is shown that the bound  $S_{m+1,n}(g)$  is better than  $S_{m,n}(g)$ , i.e.,

$$|\text{Var } g(X) - S_{m+1,n}(g)| \leq |\text{Var } g(X) - S_{m,n}(g)|.$$

Also, for any suitable function  $g$ ,  $S_{m,n}(g) \rightarrow \text{Var } g(X)$  as  $m \rightarrow \infty$ .

## 2 Universal variance bounds

This Section presents a wide class of variance bounds. First we give the following definitions that will be used in the sequel.

**Definition 2.1** For  $x \in \mathbb{R}$  and  $k \in \mathbb{N}$  define:

- (a)  $(x)_k = x(x-1) \cdots (x-k+1)$ , with  $(x)_0 = 1$ .
- (b)  $[x]_k = x(x+1) \cdots (x+k-1)$ , with  $[x]_0 = 1$ .

Note that  $[x]_k = (-1)^k (-x)_k = (x+k-1)_k$ .

**Definition 2.2** (cf. [2]) Assume that  $X \sim \text{IP}(\mu; q)$  and denote by  $q(x) = \delta x^2 + \beta x + \gamma$  its quadratic polynomial. Let  $J(X) = (\alpha; \omega)$  be the support of  $X$  and fix the non-negative integers  $m, n$  such that  $\mathbb{E}|X|^{2\ell}$  is finite, where  $\ell = \max\{m, n\}$ . We shall denote by  $\mathcal{H}^{m,n}(X)$  the class of Borel functions  $g : (\alpha, \omega) \rightarrow \mathbb{R}$  satisfying the following properties:

$H_1$  :  $g \in C^{\ell-1}(\alpha, \omega)$  and the function  $g^{(\ell-1)}(x) := \frac{d^{\ell-1}g(x)}{dx^{\ell-1}}$  is absolutely continuous in  $(\alpha, \omega)$  with a.s. derivative  $g^{(\ell)}$ .

$H_2$  :  $\mathbb{E}q^n(X)(g^{(n)}(X))^2 < \infty$  and  $\mathbb{E}q^i(X)|g^{(i)}(X)| < \infty$  for  $i = n+1, \dots, m$ , if any.

Note that for  $m = n = 0$ , the property  $H_1$  does not impose any restriction on  $g$ , and

$$\mathcal{H}^{0,0}(X) \equiv L^2(\mathbb{R}, X) := \{g : (\alpha, \omega) \rightarrow \mathbb{R} \text{ such that } \text{Var } g(X) < \infty\}.$$

Also, it is obvious that  $\mathcal{H}^{0,n} = \mathcal{H}^{1,n} = \dots = \mathcal{H}^{n,n}$ . Furthermore, we shall denote by  $\mathcal{H}^{\infty,n}(X)$  and  $\mathcal{H}^\infty(X) \equiv \mathcal{H}^{m,\infty}(X)$  [ $m$  is arbitrary because in this case this index is insignificant] the classes of functions  $\mathcal{H}^{\infty,n}(X) := \bigcap_{m=0}^\infty \mathcal{H}^{m,n}(X)$  and  $\mathcal{H}^\infty(X) := \bigcap_{n=0}^\infty \mathcal{H}^{\infty,n}(X)$ ; i.e.,

$$\mathcal{H}^{\infty,n}(X) = \{g \in C^\infty(J) : \mathbb{E}q^n(X)(g^{(n)}(X))^2 < \infty \text{ and } \mathbb{E}q^i(X)|g^{(i)}(X)| < \infty \forall i > n\},$$

$$\mathcal{H}^\infty(X) = \{g \in C^\infty(J) : \mathbb{E}q^n(X)(g^{(n)}(X))^2 < \infty \forall n \in \mathbb{N}\}.$$

From (A.3) and from  $\mathbb{E}^2 q^i(X)|g^{(i)}(X)| \leq \mathbb{E}q^i(X) \cdot \mathbb{E}q^i(X)(g^{(i)}(X))^2 < \infty$ ,  $i = 0, 1, \dots, n$ , we conclude that the (finite or infinite) sequence  $\mathcal{H}^{m,n}(X)$  is decreasing in  $m$  and in  $n$ . In particular, if all moments exist then

$$\begin{aligned} L^2(\mathbb{R}, X) &\equiv \mathcal{H}^{0,0}(X) \\ &\quad \bigcup \mathcal{H}^{1,0}(X) \supseteq \mathcal{H}^{1,1}(X) \\ &\quad \bigcup \mathcal{H}^{2,0}(X) \supseteq \mathcal{H}^{2,1}(X) \supseteq \mathcal{H}^{2,2}(X) \\ &\quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ &\quad \bigcup \mathcal{H}^{\infty,0}(X) \supseteq \mathcal{H}^{\infty,1}(X) \supseteq \mathcal{H}^{\infty,2}(X) \supseteq \dots \supseteq \mathcal{H}^\infty(X) \end{aligned}$$

Let  $X \sim \text{IP}(\mu; \delta, \beta, \gamma)$  with  $\delta \leq 0$ . Also, consider two (fixed) non-negative integers  $m, n$ , with  $n \neq 0$ , and a function  $g \in \mathcal{H}^{m,n}(X)$ . According to Parseval's identity we have that

$$\text{Var } g(X) = \sum_{k=1}^{\infty} c_k^2, \tag{2.1}$$

where  $c_k = \mathbb{E}g(X)\phi_k(X)$  are the Fourier coefficients of  $g$  with respect to the corresponding (to  $X$ ) orthonormal polynomial system  $\{\phi_k\}_{k=0}^\infty$ .

For each  $i = 1, 2, \dots, n$  the function  $g^{(i)} \in \mathcal{H}^{m-i, n-i}(X_i)$  and, from Parseval's identity again,  $\mathbb{E}(g^{(i)}(X_i))^2 = \sum_{k=0}^\infty (c_k^{(i)})^2$ , where  $c_k^{(i)} = \mathbb{E}g^{(i)}(X_i)\phi_{k,i}(X_i)$  are the Fourier coefficients of  $g^{(i)}$

with respect to the orthonormal polynomial system  $\{\phi_{k,i}\}_{k=0}^{\infty}$  corresponding to  $X_i \sim f_i \propto q^i f$ ; see Appendix A. Using (A.4) we have that

$$\mathbb{E}q^i(X)(g^{(i)}(X))^2 = \sum_{k=i}^{\infty} \left( (k)_i \prod_{j=k-1}^{k+i-2} (1-j\delta) \right) c_k^2, \quad i = 1, 2, \dots, n, \quad (2.2)$$

where each coefficient of  $c_k^2$  is positive. Also, from (A.2),

$$\mathbb{E}q^i(X)g^{(i)}(X) = \left( i! \mathbb{E}q^i(X) \prod_{j=i-1}^{2i-2} (1-j\delta) \right)^{1/2} c_i, \quad i = 1, 2, \dots, m. \quad (2.3)$$

Let  $\lambda_n = (\lambda_{1,n}, \lambda_{2,n}, \dots, \lambda_{n,n})^t \in \mathbb{R}^n$ . According to Tonelli's Theorem we have that  $\sum_{i=1}^n \sum_{k=i}^{\infty} |\lambda_{i,n}(k)_i \prod_{j=k-1}^{k+i-2} (1-j\delta)| c_k^2 = \sum_{i=1}^n |\lambda_{i,n}| \sum_{k=i}^{\infty} [(k)_i \prod_{j=k-1}^{k+i-2} (1-j\delta)] c_k^2 = \sum_{i=1}^n |\lambda_{i,n}| \times \mathbb{E}q^i(X)(g^{(i)}(X))^2 < \infty$  and, using Fubini's Theorem,

$$\sum_{i=1}^n \lambda_{i,n} \mathbb{E}q^i(X)(g^{(i)}(X))^2 = \sum_{k=1}^{\infty} \rho_{k;n} c_k^2, \quad \text{where } \rho_{k;n} = \sum_{i=1}^{\min\{k,n\}} \lambda_{i,n}(k)_i \prod_{j=k-1}^{k+i-2} (1-j\delta). \quad (2.4)$$

We seek for a vector  $\lambda_{m,n}$  such that  $\rho_{m+1,n} = \rho_{m+2,n} = \dots = \rho_{m+n,n} = 1$ . From (2.2) we obtain the system of equations

$$A_{m,n} \cdot \lambda_{m,n} = \mathbf{1}_n, \quad (2.5)$$

where the matrix  $A_{m,n} \in \mathbb{R}^{n \times n}$  has  $(r, c)$ -element which is given by

$$a_{r,c;m,n} = (m+r)_c \prod_{j=r-1}^{m+r+c-2} (1-j\delta)$$

and the vector  $\mathbf{1}_n = (1, 1, \dots, 1)^t \in \mathbb{R}^n$ .

The above system has the unique solution, see Appendix B,

$$\lambda_{i,m,n} = (-1)^{i-1} \binom{n}{i} / [(m+n)_i \prod_{j=m}^{m+i-1} (1-j\delta)], \quad i = 1, 2, \dots, n. \quad (2.6)$$

From (2.4) and (2.6), using the hypergeometric series (C.1), we have that  $\rho_{k;m,n} = 1 - (m+n-k)_n \prod_{j=m+k}^{m+n+k-1} (1-j\delta) / [(m+n)_n \prod_{j=m}^{m+n-1} (1-j\delta)]$ . Equivalently,

$$\rho_{k;m,n} = \begin{cases} 1 - \frac{(m+n-k)_n \prod_{j=m+k}^{m+n+k-1} (1-j\delta)}{(m+n)_n \prod_{j=m}^{m+n-1} (1-j\delta)}, & 1 \leq k \leq m, \\ 1 & m < k \leq m+n, \\ 1 + (-1)^{n-1} \frac{(k-m-1)_n \prod_{j=m+k}^{m+n+k-1} (1-j\delta)}{(m+n)_n \prod_{j=m}^{m+n-1} (1-j\delta)}, & k > m+n. \end{cases}$$

Thus,

$$\begin{aligned} & \sum_{i=1}^n (-1)^{i-1} \frac{\binom{n}{i} \mathbb{E}q^i(X)(g^{(i)}(X))^2}{(m+n)_i \prod_{j=m}^{m+i-1} (1-j\delta)} \\ &= \text{Var } g(X) - \sum_{k=1}^n \frac{(m+n-k)_n \prod_{j=m+k}^{m+n+k-1} (1-j\delta)}{(m+n)_n \prod_{j=m}^{m+n-1} (1-j\delta)} c_k^2 \\ & \quad + \sum_{k>m+n} (-1)^{n-1} \frac{(k-m-1)_n \prod_{j=m+k}^{m+n+k-1} (1-j\delta)}{(m+n)_n \prod_{j=m}^{m+n-1} (1-j\delta)} c_k^2. \end{aligned} \quad (2.7)$$

The main result of this paper is contained in the following theorem.

**Theorem 2.1** *Let  $X \sim \text{IP}(\mu; \delta, \beta, \gamma)$  with  $\delta \leq 0$ . Fixed two non-negative integers  $m, n$  [with  $n \neq 0$ ] and a function  $g \in \mathcal{H}^{m,n}(X)$ . Consider the quantity*

$$S_{m,n}(g) = \sum_{i=1}^m a_i \mathbb{E}^2 q^i(X) g^{(i)}(X) + \sum_{i=1}^n (-1)^{i-1} b_i \mathbb{E} q^i(X) (g^{(i)}(X))^2, \quad (2.8)$$

where

$$a_i := \frac{\binom{m}{i} \prod_{j=m+i}^{m+n+i-1} (1-j\delta)}{(m+n)_i \mathbb{E} q^i(X) (\prod_{j=i-1}^{2i-2} (1-j\delta)) \prod_{j=m}^{m+n-1} (1-j\delta)},$$

$$b_i := \frac{\binom{n}{i}}{(m+n)_i \prod_{j=m}^{m+i-1} (1-j\delta)}$$

are strictly positive constants (depending only on  $m, n$  and  $X$ ) and the empty sums (when  $m = 0$ ) are treated as zero. Then the following inequality holds:

$$(-1)^n (\text{Var } g(X) - S_{m,n}(g)) \geq 0,$$

and where  $S_{m,n}(g)$  becomes equal to  $\text{Var } g(X)$  if and only if  $g$  is a polynomial of degree at most  $m+n$ .

*Proof* From (2.8), via (2.3) and (2.7), we obtain that  $(-1)^n (\text{Var } g(X) - S_{m,n}(g)) = R_{m,n}(g)$ , where the residual

$$R_{m,n}(g) = \sum_{k>m+n} r_{k,m,n}(X) c_k^2 := \sum_{k>m+n} \frac{(k-m-1)_n \prod_{j=m+k}^{m+n+k-1} (1-j\delta)}{(m+n)_n \prod_{j=m}^{m+n-1} (1-j\delta)} c_k^2 \quad (2.9)$$

is non-negative and equals to zero if and only if  $c_k = 0$  for all  $k > m+n$ , i.e., the function  $g : J(X) \rightarrow \mathbb{R}$  is a polynomial of degree at most  $m+n$ .  $\square$

**Remark 2.1** (a) For fixed  $n$  and for any function  $g \in \mathcal{H}^{M,n}(X)$ , where  $M$  can be finite or infinite, the variance bounds  $\{S_{m,n}(g)\}_{m=0}^M$  are of the same type, i.e. upper bound when  $n$  is odd and lower bound when  $n$  is even.

(b) The bounds  $\{S_{m,n}(g)\}_{m=0}^n$  require the same conditions on the function  $g$ , i.e.,  $g \in \mathcal{H}^{n,n}(X)$ .

**Remark 2.2** (a) The bound  $S_{1,1}(g)$  is the bound  $S_{1,(\text{str})}$  of (1.4).

(b) The bounds  $S_{0,n}(g)$  are the bounds  $S_n$  which are given by (1.2). Also, for the special case  $m = 0, n = 1$  observe that  $S_{0,1}(g) = S_1 = S_{1,(\text{weak})}$ ; see (1.5).

(c) The results shown in Theorem 2.1 apply to the special case where  $n = 0$  (note that the second sum is empty and is treated as zero). In this case the lower bound  $S_{m,0}(g)$  is reduced to the one given by (1.3).

**Remark 2.3** Regarding the conditions on the function  $g$  of Theorem 2.1 we note that  $g \in \mathcal{H}^{\max\{m,n\}, n-1}(X) \setminus \mathcal{H}^{\max\{m,n\}, n}(X)$  implies that the bound  $S_{m,n}(g)$  is trivial, i.e.,  $+\infty$  when  $n$  is odd and  $-\infty$  when  $n$  is even.

Table 2.1: Specific form of the bounds  $S_{m,n}(g)$  for normal, gamma and beta distributions.

distribution	parameters	$f$	$J(X)$	$q(x)$
bounds $S_{m,n}(g)$				
normal( $\mu, \sigma^2$ )	$\mu \in \mathbb{R}, \sigma^2 > 0$	$\frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}$	$\mathbb{R}$	$\sigma^2$
gamma( $\alpha, \theta$ )	$\alpha, \theta > 0$	$\sum_{i=1}^m \frac{\binom{m}{i} \sigma^{2i}}{(m+n)_i} \mathbb{E}^2 g^{(i)}(X) + \sum_{i=1}^n (-1)^{i-1} \frac{\binom{n}{i} \sigma^{2i}}{(m+n)_i} \mathbb{E} (g^{(i)}(X))^2$	$(0, \infty)$	$x/\theta$
		$\frac{\theta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\theta x}$		
beta( $\alpha, \beta$ )	$\alpha, \beta > 0$	$\sum_{i=1}^m \frac{\binom{m}{i}}{(m+n)_i \alpha_i} \mathbb{E}^2 X^i g^{(i)}(X) + \sum_{i=1}^n (-1)^{i-1} \frac{\binom{n}{i}}{(m+n)_i \theta^i} \mathbb{E} X^i (g^{(i)}(X))^2$	$(0, 1)$	$x(1-x)/(\alpha+\beta)$
		$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$		
		$\sum_{i=1}^m \frac{\binom{m}{i} [\alpha+\beta+m+i]_m [\alpha+\beta]_{2i}}{(m+n)_i \alpha_i [\beta]_i [\alpha+\beta+i-1]_i [\alpha+\beta+m]_n} \mathbb{E}^2 X^i (1-X)^i g^{(i)}(X)$ $+ \sum_{i=1}^n (-1)^{i-1} \frac{\binom{n}{i}}{(m+n)_i [\alpha+\beta+m]_i} \mathbb{E} X^i (1-X)^i (g^{(i)}(X))^2$		

Now, we seek for upper bounds of the non-negative residual  $R_{m,n}(g)$ .

**Proposition 2.1** *Assume the conditions of Theorem 2.1 and, further, suppose that  $g \in \mathcal{H}^{T,T}(X)$  for some  $T \in \{n, \dots, m+n+1\}$ . Then the residual  $R_{m,n}(g)$ , given by (2.9), is bounded above by*

$$u_\tau \mathbb{E} q^\tau(X) (g^{(\tau)}(X))^2, \quad \tau = n, n+1, \dots, T, \quad (2.10)$$

where  $u_\tau = u_{m,n,\tau}(X) := \prod_{j=2m+n+1}^{2m+2n} (1-j\delta) / \left[ \binom{m+n}{n} (m+n+1)_\tau \prod_{j=m}^{m+n+\tau-1} (1-j\delta) \right]$ .

*Proof* Using (2.2) we write the quantity (2.10) in the form  $\sum_{k=\tau}^\infty \pi_{k;\tau} c_k^2$ . Next, consider the sequence  $\{w_{k;\tau} = \pi_{k;\tau} / r_{k;m,n}(X)\}_{k=m+n+1}^\infty$ , where  $r_{k;m,n}(X)$  are the numbers given by (2.9), and observe that this sequence is increasing in  $k$ , with  $w_{m+n+1;\tau} = 1$ .  $\square$

The upper bounds (when there are at least two) of the residual  $R_{m,n}(g)$ , given by (2.10), are not comparable. For example, consider the functions  $g_1 = \varphi_\tau$  and  $g_2 = \varphi_{m+n+2}$  (both belong to  $\mathcal{H}^\infty(X)$ ), where  $\varphi_k$  are the polynomials given by (A.1), and observe that

$$u_\tau \mathbb{E} q^\tau(X) (g_1^{(\tau)}(X)) = u_\tau \tau! \prod_{j=k-2}^{2\tau-2} (1-j\delta) > 0 = u_{\tau+1} \mathbb{E} q^{\tau+1}(X) (g_1^{(\tau+1)}(X))$$

and

$$\frac{u_\tau \mathbb{E} q^\tau(X) (g_2^{(\tau)}(X))}{u_{\tau+1} \mathbb{E} q^{\tau+1}(X) (g_2^{(\tau+1)}(X))} = \frac{(m+n-\tau+1)(1-(m+n+\tau)\delta)}{(m+n-\tau+2)(1-(m+n+\tau+1)\delta)} < 1,$$

since  $\delta \leq 0$ .

Let  $X \sim \text{IP}(\mu; \delta, \beta, \gamma)$  with  $\delta \leq 0$ . Then  $X$  is a linear function of a normal, a gamma or a beta random variable, see [2]. The bounds  $S_{m,n}(g)$  for the three main cases are included in Table 2.1.

### 3 Investigating the bounds $S_{m,n}$ for fixed $n$

Next, for  $n$  fixed, we investigate the bounds  $S_{m,n}(g)$  as  $m$  increases. We compare the variance bounds  $S_{n,n}(g)$  and  $S_n$ , given by (2.8) [for  $m=n$ ] and (1.2), respectively. Also, we compare the new upper variance bounds  $S_{n,1}(g)$  and  $S_{n-1,1}(g)$  with the existing Chernoff-type upper variance bounds  $S_{n,(\text{str})}$  and  $S_{n,(\text{weak})}$ , respectively; see (1.4) and (1.5).

**Theorem 3.1** Let  $X \sim \text{IP}(\mu; \delta, \beta, \gamma)$  with  $\delta \leq 0$ . Fix the positive integer  $n$  and consider a function  $g \in \mathcal{H}^{M,n}(X)$ , where  $M$  can be finite ( $M \geq n$ ) or infinite. Then for each  $m_1, m_2$  such that  $0 \leq m_1 < m_2 \leq M$  the following inequality holds

$$|\text{Var} g(X) - S_{m_1,n}(g)| \geq \zeta_{m_1,m_2,n}(\delta) |\text{Var} g(X) - S_{m_2,n}(g)|, \quad (3.1)$$

where  $\zeta_{m_1,m_2,n}(\delta) := \frac{(m_2+n)_n \prod_{j=m_2}^{m_2+n-1} (1-j\delta)}{(m_1+n)_n \prod_{j=m_1}^{m_1+n-1} (1-j\delta)} > 1$ . The equality holds if and only if the function  $g : J(X) \rightarrow \mathbb{R}$  is a polynomial of degree at most  $n + m_1$ .

*Proof* Consider the positive sequence  $\{\zeta_{k,m_1,m_2,n}(\delta) = r_{k,m_1,n}(X)/r_{k,m_2,n}(X)\}_{k>m_2+n}$ , where  $r_{k,m,n}(X)$  are given by (2.9). This sequence is decreasing in  $k$ . Specifically,

$$\zeta_{k,m_1,m_2,n}(\delta) \searrow \zeta_{m_1,m_2,n}(\delta) \equiv \frac{(m_2+n)_n \prod_{j=m_2}^{m_2+n-1} (1-j\delta)}{(m_1+n)_n \prod_{j=m_1}^{m_1+n-1} (1-j\delta)}, \quad \text{as } k \rightarrow \infty.$$

Moreover, we observe that  $r_{k,m_1,n}(X) > 0$  and  $r_{k,m_2,n}(X) = 0$  for all  $k = n + m_1 + 1, \dots, n + m_2$ . Therefore (3.1) follows.

We write  $|\text{Var} g(X) - S_{m_1,n}(g)| - \zeta_{m_1,m_2,n}(\delta) |\text{Var} g(X) - S_{m_2,n}(g)| = \sum_{k>n+m_1} \theta_k c_k^2$  and we observe that  $\theta_k > 0$  for all  $k$ . Thus, the equality in (3.1) holds if and only if  $g$  is a polynomial of degree at most  $n + m_1$ .  $\square$

Notice that if  $\delta < 0$  then  $\prod_{j=m_2}^{m_2+n-1} (1-j\delta) / \prod_{j=m_1}^{m_1+n-1} (1-j\delta) > 1$  for each  $n$  and  $m_1 < m_2$ . Therefore,  $\zeta_{m_1,m_2,n}(\delta) \geq \zeta_{m_1,m_2,n}(0) = (m_2+n)_n / (m_1+n)_n$ , since  $\delta \leq 0$ .

**Remark 3.1** Assume the conditions of Theorem 3.1. (a) In view of Remark 2.1(a), the bounds  $\{S_{m,n}(g)\}_{m=0}^M$  are of the same type. From (3.1) it follows that the bound  $S_{m_2,n}(g)$  is better than the bound  $S_{m_1,n}(g)$ . Now, writing  $n = 2r$  (when  $n$  is even) or  $n = 2r + 1$  (when  $n$  is odd) we observe that

$$S_{0,2r}(g) \leq S_{1,2r}(g) \leq \dots \leq \text{Var} g(X) \leq \dots \leq S_{1,2r+1}(g) \leq S_{0,2r+1}(g).$$

(b) For the case  $M = \infty$ , from (2.1), (2.8) and (a) it follows that

$$S_{m,n}(g) \nearrow \text{Var} g(X) \text{ or } S_{m,n}(g) \searrow \text{Var} g(X), \quad \text{as } m \rightarrow \infty.$$

[when  $n$  is even]                      [when  $n$  is odd]

Now, we compare the existing variance bounds  $S_n(\equiv S_{0,n}(g))$  with the best proposed bound shown in this article requiring the same conditions on  $g$ , i.e., with the bound  $S_{n,n}(g)$ ; see Remark 2.1(b).

**Corollary 3.1** The variance bounds  $S_{n,n}(g)$  and  $S_n$ , given by (2.8) (for  $m = n$ ) and (1.2) respectively, are of the same type and require the same conditions on  $g$ . Moreover, the new bound  $S_{n,n}(g)$  is better than the existing bound  $S_n$ . Specifically,

$$|\text{Var} g(X) - S_n| \geq \binom{2n}{n} \frac{\prod_{j=n}^{2n-1} (1-j\delta)}{\prod_{j=0}^{n-1} (1-j\delta)} |\text{Var} g(X) - S_{n,n}(g)|.$$

The equality holds only in the trivial cases when  $\text{Var} g(X) = S_{n,n}(g) = S_n$ , i.e., the function  $g : J(X) \rightarrow \mathbb{R}$  is a polynomial of degree at most  $n$ .

Note that, since  $\delta \leq 0$ ,  $\binom{2n}{n} \prod_{j=n}^{2n-1} (1-j\delta) / \prod_{j=0}^{n-1} (1-j\delta) \geq \binom{2n}{n}$ .

The quantities  $S_{n,(\text{str})}$  and  $S_{n,1}(g)$  are upper bounds for  $\text{Var}g(X)$ . Both bounds are equal to  $\text{Var}g(X)$  if and only if the function  $g$  is a polynomial of degree at most  $n+1$ . The quantities  $S_{n,(\text{weak})}$  and  $S_{n-1,1}(g)$  are upper bounds for  $\text{Var}g(X)$ . Both bounds are equal to  $\text{Var}g(X)$  if and only if the function  $g$  is a polynomial of degree at most  $n$ . Thus, it is reasonable to compare these bounds.

**Theorem 3.2** *For  $n = 1, 2, \dots$  and any suitable function  $g$  we have that:*

(a)  $S_{n,1}(g) \leq S_{n,(\text{str})}$ , where the equality holds when  $n = 1$  or  $n > 1$  and  $g$  is a polynomial of degree at most  $n+1$ .

(b)  $S_{n-1,1}(g) \leq S_{n,(\text{weak})}$ , where the equality holds when  $n = 1$  or  $n > 1$  and  $g$  is a polynomial of degree at most  $n$ .

*Proof* (a) From (2.8) and (1.4), via (2.2) and (2.3), we have that

$$S_{n,(\text{str})} - S_{n,1}(g) = \sum_{k>n+1} \frac{k[1-(k-1)\delta]}{(n+1)(1-n\delta)} \left( \frac{\binom{k-1}{n-1} \prod_{j=k}^{n+k-2} (1-j\delta)}{n \prod_{j=n+1}^{2n-1} (1-j\delta)} - 1 \right) c_k^2, \quad (3.2)$$

where each coefficient of  $c_k^2$ ,  $k > n+1$ , is zero when  $n = 1$  and is positive when  $n > 1$ .

(b) Similarly, from (2.8) and (1.5), via (2.2) and (2.3), it follows that

$$S_{n,(\text{weak})} - S_{n-1,1}(g) = \sum_{k>n} \frac{k[1-(k-1)\delta]}{n[1-(n-1)\delta]} \left( \frac{\binom{k-1}{n-1} \prod_{j=k}^{n+k-2} (1-j\delta)}{\prod_{j=n}^{2n-2} (1-j\delta)} - 1 \right) c_k^2, \quad (3.3)$$

where each coefficient of  $c_k^2$ ,  $k > n$ , is zero when  $n = 1$  and is positive when  $n > 1$ .  $\square$

**Remark 3.2** For each  $n = 2, 3, \dots$  it follows that:

(a) The bound  $S_{n,1}(g)$  is better than the bound  $S_{n,(\text{str})}$ ; notice that the bound  $S_{n,1}(g)$  requires a milder finiteness condition,  $g \in \mathcal{H}^{n,1}(X)$ , compared to  $S_{n,(\text{str})}$  which requires that  $g \in \mathcal{H}^{n,n}(X) \subseteq \mathcal{H}^{n,1}(X)$ .

(b) The bound  $S_{n-1,1}(g)$  is better than the bound  $S_{n,(\text{weak})}$ ; as in (a), the bound  $S_{n-1,1}(g)$  requires a weaker finiteness condition, i.e.  $g \in \mathcal{H}^{n-1,1}(X)$ , rather than  $S_{n,(\text{weak})}$ , i.e.  $g \in \mathcal{H}^{n,n}(X) \subseteq \mathcal{H}^{n-1,1}(X)$ .

**Final Conclusion** *The variance bounds given by Theorem 2.1, for appropriate choices of  $n$  and  $m$ , either provide existing univariate variance bounds or improvements. Our bounds cover all usual cases, namely:*

- Chernoff-type, see [13, 3, 8, 5, 14, 17, 2];
- Poincaré-type, see [15, 11, 10, 1];
- Bessel-type, see [4, 1].

*Notice that no further conditions on the function  $g$  are imposed; instead the new bounds require the same or weaker conditions, see Remarks 2.1, 2.2 and 3.2, Theorem 3.2, and Corollary 3.1. Therefore, the new proposed variance bounds outweigh all the existing variance bounds presented in the bibliography.*



## Appendix

### A Some useful properties of the Integrated Pearson Family

Consider a random variable  $X$  with density  $f \sim \text{IP}(\mu; q) \equiv \text{IP}(\mu; \delta, \beta, \gamma)$ .

Let  $a \in (1, +\infty)$ . Then  $E|X|^a < \infty$  if and only if  $\delta < 1/(a-1)$ . Notice that  $X$  has finite moments of any order if and only if  $\delta \leq 0$ ; see [2, Corollary A.2].

The support of  $X$  is the interval  $J(X) = (\alpha, \omega)$  and the density  $f \in C^\infty(\alpha, \omega)$ , see [2].

If  $E|X|^{2i+1} < \infty$ ,  $i \in \mathbb{N}^* \equiv \mathbb{N} \setminus \{0\}$  (that is,  $\delta < 1/2i$ ), then the random variable  $X_i$  with density function  $f_i(x) = q^i(x)f(x)/\mathbb{E}q^i(X)$  follows  $\text{IP}(\mu_i; q_i)$  distribution with  $\mu_i = (\mu + i\beta)/(1 - 2i\delta)$  and  $q_i = q/(1 - 2i\delta)$ ; see [2, Theorem D.2]. Note that if  $E|X|^{2i} < \infty$  and  $E|X|^{2i+1} = \infty$  then the function  $f_i$  is a probability density function; however,  $E|X_i| = \infty$  so  $X_i$  does not belong to the Integrated Pearson Family.

If  $E|X|^{2N} < \infty$ ,  $N \in \mathbb{N}^*$  (that is,  $\delta < 1/(2N-1)$ ), then the quadratic  $q$  generates the orthogonal polynomials through the Rodrigues-type formula, [16],

$$P_k(x) = \frac{(-1)^k}{f(x)} \frac{d^k}{dx^k} [q^k(x)f(x)], \quad x \in J(X), \quad k = 0, 1, \dots, N.$$

Afendras et al., [1], showed an extended Stein-type identity of order  $n$ . This identity takes the form  $\mathbb{E}P_k(X)g(X) = \mathbb{E}q^k(X)g^{(k)}(X)$ , provided that  $\mathbb{E}q^k(X)|g^{(k)}(X)| < \infty$ . Also,  $\mathbb{E}P_k(X)P_m(X) = \delta_{k,m}k!\mathbb{E}q^k(X)\prod_{j=k-1}^{2k-2}(1-j\delta)$ ; thus, the system of polynomials  $\{\varphi_k\}_{k=0}^N$  is orthonormal, with respect to density  $f$ , where

$$\varphi_k(x) = P_k(x) \left( k! \mathbb{E}q^k(X) \prod_{j=k-1}^{2k-2} (1-j\delta) \right)^{-1/2}. \quad (\text{A.1})$$

Hence,

$$\mathbb{E}q^k(X)g^{(k)}(X) = \left( k! \mathbb{E}q^k(X) \prod_{j=k-1}^{2k-2} (1-j\delta) \right)^{-1/2} \mathbb{E}\varphi_k(X)g(X). \quad (\text{A.2})$$

Moreover, the system of polynomials  $\{\varphi_{k+i}^{(i)}\}_{k=0}^{N-i}$  [where  $\varphi_k$  are the polynomials given by (A.1) and  $\varphi_{k+i}^{(i)}$  is the  $i$ -th derivative of  $\varphi_{k+i}$ ] is orthogonal with respect to density  $f_i$ . Specifically, if  $\varphi_{k,i}$  are the orthonormal polynomials corresponding to the density  $f_i$  then  $\varphi_{k+i}^{(i)}(x) = v_k^{(i)} \varphi_{k,i}(x)$ , where  $v_k^{(i)} = v_k^{(i)}(X) := [(k+i)_i (\prod_{j=k+i-1}^{k+2i-2} (1-j\delta)) / \mathbb{E}q^i(X)]^{1/2}$ , see [2, Corollary D.4].

**Lemma A.1** (see [2]) *Let a random variable  $X \sim \text{IP}(\mu, q)$  and consider the strictly positive integers  $n$  and  $N$  such that  $n \leq N$  and  $E|X|^{2N} < \infty$ .*

$$\text{If } g \in \mathcal{H}^{n,n}(X) \text{ then } g^{(i)} \in \mathcal{H}^{n-i,n-i}(X_i) \text{ for each } i = 0, 1, \dots, n-1. \quad (\text{A.3})$$

$$\mathbb{E}\varphi_{k,i}(X_i)g^{(i)}(X_i) = v_k^{(i)}(X)\mathbb{E}\varphi_{k+i}(X)g(X) \text{ for each } \begin{cases} i = 1, 2, \dots, n, \\ k = 0, 1, \dots, N-i. \end{cases} \quad (\text{A.4})$$

If the parameter  $\delta$  of  $q$  is non-positive then the moment generating function of  $X$  is finite in a neighborhood of zero; thus the system of polynomials  $\{\varphi_k\}_{k=0}^\infty$  forms an orthonormal basis of  $L^2(\mathbb{R}, X)$  and the Parseval identity holds, see [2]. Notice that for each  $i \in \mathbb{N}$  the parameter  $\delta_i = \frac{\delta}{1-2i\delta}$  is non-positive too. Thus, the system of polynomials  $\{\varphi_{k,i}\}_{k=0}^\infty$  is an orthonormal basis of  $L^2(\mathbb{R}, X_i)$  and the Parseval identity holds too.

### B The solution of the system (2.5)

Consider the determinants  $d_{m,n} = \det(A_{m,n})$  and  $d_{i,m,n} = \det(A_{i,m,n})$ ,  $i = 1, 2, \dots, n$ , where the matrix  $A_{i,m,n}$  is formed from  $A_{m,n}$  by replacing column  $i$  with the vector  $\mathbf{1}_n$ .

For each  $t = 1, 2, \dots, n$  define the matrix  $B_{m,n}(t) \in \mathbb{R}^{(n-t+1) \times (n-t+1)}$  [ $m, n$  are fixed] which has elements  $b_{r,c;m,n}(t) = (m+r-1)_{c-1} \prod_{j=m+r+t-1}^{m+r+c+t-3} (1-j\delta)$ , where empty products are treated as one. Observe that

$d_{m,n} = (m+n)_n (\prod_{j=m}^{m+n-1} (1-j\delta)) \det(B_{m,n}(1))$  and  
 $\det(B_{m,n}(t)) = (n-t)! (\prod_{j=m+1}^{m+n-t} (1-[2j+(t-1)]\delta)) \det(B_{m,n}(t+1)), t = 1, 2, \dots, n-1.$

Thus, it follows that

$$d_{m,n} = (m+n)_n [\prod_{j=0}^{n-1} j!] [\prod_{j=m}^{m+n-1} (1-j\delta)] \prod_{t=1}^{n-1} (\prod_{j=m+1}^{m+n-t} (1-[2j+(t-1)]\delta)) \neq 0.$$

Now, for each  $t = 1, 2, \dots, n$  define the matrix  $B_{i,m,n}(t) \in \mathbb{R}^{(n-t+1) \times (n-t+1)}$  [ $i, m, n$  are fixed integers] which has  $(r, c)$ -element  $b_{r,c;i,m,n}(t) = (m+r)_{c-1} \prod_{j=m+r+t-2}^{m+r+c+t-4} (1-j\delta)$ , when  $c = 1, 2, \dots, i-t$ , and  $b_{r,c;i,m,n}(t) = (m+r)_c \prod_{j=m+r+t-2}^{m+r+c+t-3} (1-j\delta)$ , when  $c = i-t+1, i-t+2, \dots, n-t+1$ . Observe that

$$d_{i,m,n} = (-1)^{i-1} \det(B_{i,m,n}(1)),$$

$$\det(B_{i,m,n}(t)) = \frac{(n-t+1)!}{(i-t+1)!} (\prod_{j=m+1}^{m+n-t} (1-[2j+(t-1)]\delta)) \det(B_{i,m,n}(t+1)), t = 1, 2, \dots, i-1, \text{ and}$$

$$\det(B_{i,m,n}(i)) = (n-i+1)! (\prod_{j=m+1}^{m+n-i} (1-[2j+(i-1)]\delta)) (m+n-i)_{n-i} (\prod_{j=m+i}^{m+n-1} (1-j\delta)) \det(B_{m,n}(i+1)).$$

Thus, it follows that

$$d_{i,m,n} = (-1)^{i-1} \frac{(m+n-i)_{n-i}}{i!(n-i)!} [\prod_{j=0}^n j!] [\prod_{j=m}^{m+n-1} (1-j\delta)] \prod_{t=1}^{n-1} (\prod_{j=m+1}^{m+n-t} (1-[2j+(t-1)]\delta)).$$

Therefore, according to Cramér's rule, (2.6) follows.

## C A useful hypergeometric series

**Lemma C.1** *Let  $m, n, k \in \mathbb{N}$  and  $\delta \leq 0$ . Then the following hypergeometric series holds:*

$$\sum_{i=0}^n (-1)^i \binom{n}{i} \frac{(k)_i}{(m+n)_i} \frac{\prod_{j=k-1}^{k+i-2} (1-j\delta)}{\prod_{j=m}^{m+i-1} (1-j\delta)} = \frac{(m+n-k)_n \prod_{j=m+k}^{m+n+k-1} (1-j\delta)}{(m+n)_n \prod_{j=m}^{m+n-1} (1-j\delta)}. \quad (\text{C.1})$$

*Proof* For the case  $\delta = 0$ , write (C.1) as  $\sum_{i=0}^n (-1)^i \binom{n}{i} \frac{(k)_i}{(m+n)_i} = \frac{(m+n-k)_n}{(m+n)_n}$  and observe that this follows from Vandermonde's formula  $\sum_{i=0}^n (-1)^i \binom{n}{i} \frac{(x)_i}{(x+y)_i} = \frac{(y)_n}{(x+y)_n}$ , by replacing  $x$  with  $k$  and  $y$  with  $m+n-k$ , see [6, p. 125].

For the case  $\delta < 0$ , write (C.1) as  $\sum_{i=0}^n (-1)^i \frac{(n)_i (k)_i (1/\delta + 1 - k)_i}{i! (m+n)_i (1/\delta - m)_i} = \frac{(m+n-k)_n (1/\delta - m - k)_n}{(m+n)_n (1/\delta - m)_n}$ . This follows from Dougall's identity,  $\sum_{i=0}^s \frac{(\alpha)_i (\beta)_i (s)_i}{i! [\gamma+1]_i (\alpha+\beta+\gamma+s)_i} = \frac{[\alpha+\gamma+1]_s [\beta+\gamma+1]_s}{[\gamma+1]_s [\alpha+\beta+\gamma+1]_s}$ , using the substitution  $\alpha \mapsto k$ ,  $\beta \mapsto (1/\delta + 1 - k)$ ,  $\gamma \mapsto (-m - n - 1)$  and  $s \mapsto n$ , see [9, eq. (2)].  $\square$

## References

- [1] AFENDRAS, G., PAPADATOS, N. and PAPATHANASIOU, V. (2011). An extended Stein-type covariance identity for the Pearson family, with applications to lower variance bounds. *Bernoulli*, **17**(2), 507–529.
- [2] AFENDRAS, G. and PAPADATOS, N. (2011). Strengthened Chernoff-type variance bounds. *Submitted for publications*. arXiv: math.PR/1107.1754
- [3] BRASCAMP, H. and LIEB, E. (1976). On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Functional Analysis*, **22**(4), 366–389.
- [4] CACOULLOS, T. (1982). On upper and lower bounds for the variance of a function of a random variable. *Ann. Probab.*, **10**, 799–809.
- [5] CACOULLOS, T. and PAPATHANASIOU, V. (1985). On upper bounds for the variance of functions of random variables. *Statist. Probab. Lett.*, **3**, 175–184.
- [6] CHARALAMBIDES, C. (2002). *Enumerative Combinatorics*. Chapman & Hall/CRC.
- [7] CHEN, L.H.Y. (1982). An inequality for the multivariate normal distribution. *J. Multivariate Anal.*, **12**, 306–315.
- [8] CHERNOFF, H. (1981). A note on an inequality involving the normal distribution. *Ann. Probab.*, **9**, 533–535.
- [9] DOUGALL, J. (1907). On Vandermonde's Theorem, and some more general Expansions. *Proc. Edinburgh Math. Soc.*, **25**, 114–132
- [10] HOUDRE, C. and KAGAN, A. (1995). Variance inequalities for functions of Gaussian variables. *J. Theoret. Probab.*, **8**, 23–30.

- [11] JOHNSON, R.W. (1993). A note on variance bounds for a function of a Pearson variate. *Statist. Decisions*, **11**, 273–278.
- [12] KORWAR, R.M. (1991). On characterizations of distributions by mean absolute deviation and variance bounds. *Ann. Inst. Statist. Math.*, **43**, 287–295.
- [13] NASH, J. (1958). Continuity of solutions of parabolic and elliptic equations *Amer. J. Math.*, **80**, 931-954.
- [14] PAPADATOS, N. and PAPATHANASIOU, V. (2001). Unified variance bounds and a Stein-type identity. In: *Probability and Statistical Models with Applications* (Ch.A. Charalambides, M.V. Koutras and N. Balakrishnan, Eds.), Chapman & Hall/CRC, New York, pp. 87-100.
- [15] PAPATHANASIOU, V. (1988). Variance bounds by a generalization of the Cauchy-Schwarz inequality. *Statist. Probab. Lett.*, **7**, 29–33.
- [16] PAPATHANASIOU, V. (1995). A characterization of the Pearson system of distributions and the associated orthogonal polynomials. *Ann. Inst. Statist. Math.*, **47**, 171–176.
- [17] PRAKASA RAO, B.L.S. (2006). Matrix variance inequalities for multivariate distributions. *Statistical Methodology*, **3**, 416–430.
- [18] WEI, Z. and ZHANG, X. (2009). Covariance matrix inequalities for functions of Beta random variables. *Statist. Probab. Lett.* **79** 873–879.